

Abstract of the Disclosure

A method for extracting Anchorable Information Units (AIUs), from a Portable Document Format (PDF) file, which
5 may either be created using either an editor or by scanning in documents. The method includes parsing the portable document format document into textual portions and non-text portions, and extracting structure from the textual portions and the non-text portions. The method further
10 includes determining text within textual portions, and text the non-text portions, and hyperlinking a plurality of keywords within the textual portions and non-text portions to a related document.